



UNITED STATES PATENT AND TRADEMARK OFFICE

UNITED STATES DEPARTMENT OF COMMERCE
United States Patent and Trademark Office
Address: COMMISSIONER FOR PATENTS
P.O. Box 1450
Alexandria, Virginia 22313-1450
www.uspto.gov

APPLICATION NO.	FILING DATE	FIRST NAMED INVENTOR	ATTORNEY DOCKET NO.	CONFIRMATION NO.
-----------------	-------------	----------------------	---------------------	------------------

10/573,736

03/28/2006

Nevenka Dimtrva

US030388

2260

24737

7590

01/22/2009

PHILIPS INTELLECTUAL PROPERTY & STANDARDS

P.O. BOX 3001

BRIARCLIFF MANOR, NY 10510

EXAMINER

ABEBE, DANIEL DEMELASH

ART UNIT

PAPER NUMBER

2626

MAIL DATE

DELIVERY MODE

01/22/2009

PAPER

Please find below and/or attached an Office communication concerning this application or proceeding.

The time period for reply, if any, is set in the attached communication.

Office Action Summary	Application No. 10/573,736	Applicant(s) DIMTRVA ET AL.	
	Examiner Daniel D. Abebe	Art Unit 2626	

-- The MAILING DATE of this communication appears on the cover sheet with the correspondence address --

Period for Reply

A SHORTENED STATUTORY PERIOD FOR REPLY IS SET TO EXPIRE 3 MONTH(S) OR THIRTY (30) DAYS, WHICHEVER IS LONGER, FROM THE MAILING DATE OF THIS COMMUNICATION.

- Extensions of time may be available under the provisions of 37 CFR 1.136(a). In no event, however, may a reply be timely filed after SIX (6) MONTHS from the mailing date of this communication.
- If NO period for reply is specified above, the maximum statutory period will apply and will expire SIX (6) MONTHS from the mailing date of this communication.
- Failure to reply within the set or extended period for reply will, by statute, cause the application to become ABANDONED (35 U.S.C. § 133). Any reply received by the Office later than three months after the mailing date of this communication, even if timely filed, may reduce any earned patent term adjustment. See 37 CFR 1.704(b).

Status

- 1) ☐ Responsive to communication(s) filed on ____.
- 2a) ☐ This action is **FINAL**. 2b) ☒ This action is non-final.
- 3) ☐ Since this application is in condition for allowance except for formal matters, prosecution as to the merits is closed in accordance with the practice under *Ex parte Quayle*, 1935 C.D. 11, 453 O.G. 213.

Disposition of Claims

- 4) ☒ Claim(s) 1-33 is/are pending in the application.
- 4a) Of the above claim(s) ____ is/are withdrawn from consideration.
- 5) ☐ Claim(s) ____ is/are allowed.
- 6) ☒ Claim(s) 1-33 is/are rejected.
- 7) ☐ Claim(s) ____ is/are objected to.
- 8) ☐ Claim(s) ____ are subject to restriction and/or election requirement.

Application Papers

- 9) ☐ The specification is objected to by the Examiner.
- 10) ☐ The drawing(s) filed on ____ is/are: a) ☐ accepted or b) ☐ objected to by the Examiner.
Applicant may not request that any objection to the drawing(s) be held in abeyance. See 37 CFR 1.85(a).
Replacement drawing sheet(s) including the correction is required if the drawing(s) is objected to. See 37 CFR 1.121(d).
- 11) ☐ The oath or declaration is objected to by the Examiner. Note the attached Office Action or form PTO-152.

Priority under 35 U.S.C. § 119

- 12) ☐ Acknowledgment is made of a claim for foreign priority under 35 U.S.C. § 119(a)-(d) or (f).
- a) ☐ All b) ☐ Some * c) ☐ None of:
1. ☐ Certified copies of the priority documents have been received.
 2. ☐ Certified copies of the priority documents have been received in Application No. ____.
 3. ☐ Copies of the certified copies of the priority documents have been received in this National Stage application from the International Bureau (PCT Rule 17.2(a)).

* See the attached detailed Office action for a list of the certified copies not received.

Attachment(s)

- | | |
|---|---|
| 1) <input checked="" type="checkbox"/> Notice of References Cited (PTO-892) | 4) <input type="checkbox"/> Interview Summary (PTO-413)
Paper No(s)/Mail Date. ____. |
| 2) <input type="checkbox"/> Notice of Draftsperson's Patent Drawing Review (PTO-948) | 5) <input type="checkbox"/> Notice of Informal Patent Application |
| 3) <input checked="" type="checkbox"/> Information Disclosure Statement(s) (PTO/SB/08)
Paper No(s)/Mail Date ____. | 6) <input type="checkbox"/> Other: ____. |

Claim Rejections - 35 USC § 101

35 U.S.C. 101 reads as follows:

Whoever invents or discovers any new and useful process, machine, manufacture, or composition of matter, or any new and useful improvement thereof, may obtain a patent therefor, subject to the conditions and requirements of this title.

Claims 1-11 and 23-33 are rejected under 35 U.S.C. 101.

the claimed invention is directed to non-statutory subject matter, because, the specification indicates that the apparatus comprising the content synthesis application can be software software. According to the specification on Par.0016 it reads

“the term "controller," "processor," or "apparatus" means any device, system or part thereof that controls at least one operation, such a device may be implemented in hardware, firmware or software, or some combination of at least two of the same”.

And further on Par.0033-0034 the specification reads

“ Content synthesis application software 235 comprises (1) a module 310 for obtaining the visual display of a face, (2) a module 320 for tracking facial features, (3) a learning module 330, (4) a module 340 for obtaining a speech portion of audio, (5) a module 350 for extracting audio features of speech, (6) a facial audio visual feature matching and classification module 360, (7) a facial animation for selected parameters module 370, and (8) a speaking face animation and synchronization module 380. The functions of the software modules will be described more fully below.

“ Content synthesis application processor 190 comprises controller 230 and content synthesis application software 235. **Controller 230 and content synthesis application software 235 together comprise a content synthesis application processor that is capable of carrying out the present invention**”

Art Unit: 2626

Thus reading independent claim 1 in light of the specification one can conclude that the claimed apparatus comprising the content synthesis application of claim 1 is Software that doesn't fall within one of the four category of statutory subject matter under 35 USC 101.

Appropriate correction can be made by amending the specification to specifically point that the claimed synthesis is implemented solely in hardware or combination of hardware and software or by amending the claim to include statutory subject matter.

With respect to claims 23-33, the claimed "synthesized audiovisual signal" is not patentable because it doesn't fall within one of the four categories of statutory subject matter cited under 35 USC 101, i.e, process, machine, manufacture or composition of matter .

DETAILED ACTION

Claim Rejections - 35 USC § 103

The following is a quotation of 35 U.S.C. 103(a) which forms the basis for all obviousness rejections set forth in this Office action:

(a) A patent may not be obtained though the invention is not identically disclosed or described as set forth in section 102 of this title, if the differences between the subject matter sought to be patented and the prior art are such that the subject matter as a whole would have been obvious at the time the invention was made to a person having ordinary skill in the art to which said subject matter pertains. Patentability shall not be negated by the manner in which the invention was made.

Claims 1-3, 7-14, 18-25 and 29-33 are rejected under 35 U.S.C. 103(a) as being unpatentable over Francini et al. (7,123,262) and further in view of McMillan et al. (6,661,418).

As to claim 1, Francini teaches an apparatus that is capable of receiving audio-visual input signals that represent a speaker who is speaking and capable of creating an animated version of the face of the speaker using a plurality of phonemes that represent the speaker's speech, said apparatus comprising a content synthesis application processor that:

extracts audio features of the speaker's speech and visual features of the speaker's face from the audio-visual input signals;

creates audiovisual input vectors from the audio features and the visual features;
creates audiovisual configurations from the audiovisual input vectors;

and obtaining an association between phonemes that represent the speaker's speech and visemes that represent the speaker's face (abstract; Figs. 2, 6; Col.2, line 55-Col.3, line 15; Col.8, lines 22-40; Claim 1). Francini doesn't explicitly teach performing semantic association as claimed.

McMillan, however, teaches a system for generating a realistic animated image of a character which is speaking, with the face of the character having visible articulation or expression matching the words being spoken, including the steps performing semantic association to obtain an expression between the text being spoken (phoneme) and the animated face (Figs.1-16; abstract; Col.8, lines 25-40). It would be obvious to one of ordinary skill in the art to combine the two teachings for the purpose generating a more realistic face animation that reflects the emotional expression within the words that are being spoken.

Art Unit: 2626

As to claim 2, Francini teaches wherein the content synthesis application processor is capable of analyzing an input audio signal by:

extracting audio features (phoneme) of a speaker's speech;

finding corresponding video representations (viseme) for the audio features using a mapping/correlation procedure;

and matching the corresponding video representations with the audiovisual configurations (face animation parameter) (Fig.1) and McMillan teaches where the face animation includes associating semantics.

As to claim 3, Francini teaches wherein the content synthesis application processor is further capable of: creating a computer generated animated face for each selected audiovisual configuration; synchronizing each computer generated animated face with the speaker's speech; and outputting an audio-visual representation of the speaker's face synchronized with the speaker's speech (Figs.1, 2, 6) and McMillan also teaches acoustically driven computer generated realistic animation (Figs.15-16).

As to claim 7, Francini teaches where the content synthesis application processor creates a number of facial animation parameter (FAP) that correspond to a particular facial expression/configuration during articulation of a particular speech (Fig.2).

As to claim 8, Francini teaches where the animation version is structured on a three-dimension model and where the animation is generated from a real video (Figs.1-2).

As to claim 9, McMillan shows a canonical face and mouth shape model to represent the different semantic expression (Figs.15-16; Col.19, lines 1-15).

As to claim 10, Francini teaches wherein said audiovisual configurations comprise audiovisual speaking face movement components (Fig.1).

As to claim 11 Francini and McMillan teach a speaking face animation and synchronization module that synchronizes each animated version of the face of the speaker with the audio features of the speaker's speech to create an audio-visual representation of the speaker's face that is synchronized with the speaker's speech;

and McMillan teaches wherein the face of the character having visible articulation or expression matching the words being spoken comprises determining a level of audio expression of the speaker's speech and providing said level of audio expression of the speaker's speech to said speaking face animation and synchronization module to modify animated facial parameters of the speaker (Figs.16).

according to McMillan " The expressions in the morph sequences portray the current emotion of the character, typically smiling or frowning. In addition to these simplistic expressions, more complicated sequences of expressions can be inserted (Col.20, lines 30-40)

Method claims 12-14 and 18-22, reciting the corresponding steps for synthesizing audio visual content by the same apparatus, are analogous to the apparatus claims addressed above and are rejected by Francini in view of McMillan for the foregoing reasons.

Claims 23-25 and 29-33, reciting the synthesized audio visual signal according to the method and using the same apparatus, are analogous and therefore rejected by Francini in view of McMillan for the foregoing reasons.

Claims 4-6, 15-17 and 26-28 are rejected under 35 U.S.C. 103(a) as being unpatentable over Francini et al. (7,123,262) in view of McMillan et al. (6,661,418) and further in view of Basu et al. (6,366,885).

As to claim 4, Francini and McMillan do not explicitly teach where the extracted feature comprise the claimed characters. Basu teaches a system for real time face animation using viseme based HMM model comprising the steps of extracting audio features where the extracted features comprises : Mel Cepstral Coefficients, Linear Predictive Coding Coefficients (Figs.2, 4). Extracting the claimed audio features are obvious in Francini and/or McMillan teachings for use to classify and associate the input audio signal with the corresponding video representation.

As to claims 5-6, Basu teaches wherein said content synthesis application uses HMM model to create and match the visual features with the audio visual features. (abstract; Figs.2-5; Col.2, line 60-Col.3, line 40).

The utilization of HMM model in Francini system will be obvious to one of ordinary skill in the art, in view of Basu teaching, as an alternative audiovisual model for the mapping of the audio parameter with the corresponding image parameters in order to generate the animation.

Claims 15-17 and 26-28 are analogous to claims 4-6 and are rejected by Francini in view of McMillan and further in view of Basu for the foregoing reasons.

Conclusion

The prior art made of record and not relied upon is considered pertinent to applicant's disclosure.

Pogio et al. (7,168,953) see entire document

“A method and apparatus for video realistic, speech animation is disclosed. A human subject is recorded using a video camera as he/she utters a predetermined speech corpus. After processing the corpus automatically, a visual speech module is learned from the data that is capable of synthesizing the human subject's mouth uttering entirely novel utterances that were not recorded in the original video. The synthesized utterance is re-composited onto a background sequence which contains natural head and eye movement. The final output is video realistic in the sense that it looks like a video camera recording of the subject. The two key components of this invention are 1) a multidimensional morphable model (MMM) to synthesize new, previously unseen mouth configurations from a small set of mouth image prototypes; and 2) a trajectory synthesis technique based on regularization, which is automatically trained from the recorded video corpus, and which is capable of synthesizing trajectories in MMM space corresponding to any desired utterance” (Abstract)

Any inquiry concerning this communication or earlier communications from the examiner should be directed to Daniel D. Abebe whose telephone number is 571-272-7615. The examiner can normally be reached on monday-friday.

If attempts to reach the examiner by telephone are unsuccessful, the examiner's supervisor, David Hudspeth can be reached on 571-272-7843. The fax phone number for the organization where this application or proceeding is assigned is 571-273-8300.

Art Unit: 2626

Information regarding the status of an application may be obtained from the Patent Application Information Retrieval (PAIR) system. Status information for published applications may be obtained from either Private PAIR or Public PAIR. Status information for unpublished applications is available through Private PAIR only. For more information about the PAIR system, see <http://pair-direct.uspto.gov>. Should you have questions on access to the Private PAIR system, contact the Electronic Business Center (EBC) at 866-217-9197 (toll-free). If you would like assistance from a USPTO Customer Service Representative or access to the automated information system, call 800-786-9199 (IN USA OR CANADA) or 571-272-1000.

/Daniel D Abebe/
Primary Examiner, Art Unit 2626